"U" shape with a connection between passing features at a greatest level of compression or abstractness from the encoder to the decoder. Any now known or later developed U-Net or other generator architectures may be used, such as including a densenet. Other fully convolutional networks may be used.

[0073] For applications, the generator of the GAN is used without the discriminator. The GAN is applied to the images from different modalities for a given patient by the generator without the discriminator. The discriminator is used for training.

[0074] The GAN is a deep architecture, which may include convolutional neural network (CNN) or deep belief nets (DBN). Other deep networks may be used. CNN learns feed-forward mapping functions, while DBN learns a generative model of data. In addition, CNN uses shared weights for all local regions, while DBN is a fully connected network (i.e., having different weights for all regions of an image). The training of CNN is entirely discriminative through back-propagation. DBN, on the other hand, employs the layer-wise unsupervised training (e.g., pre-training), followed by the discriminative refinement with back-propagation if necessary.

[0075] The network is defined as a plurality of sequential feature units or layers. Sequential is used to indicate the general flow of output feature values from one layer to input to a next layer. The information from the next layer is fed to a next layer, and so on until the final output. The layers may only feed forward or may be bi-directional, including some feedback to a previous layer. The nodes of each layer or unit may connect with all or only a sub-set of nodes of a previous or subsequent layer or unit.

[0076] Rather than pre-programming the features and trying to relate the features to attributes, the deep architecture is defined to learn the features at different levels of abstraction based on input images with or without pre-processing. The features are learned to reconstruct lower level features (i.e., features at a more abstract or compressed level). For example, features for reconstructing an image are learned. For a next unit, features for reconstructing the features of the previous unit are learned, providing more abstraction. Each node of the unit represents a feature. Different units are provided for learning different features.

[0077] Within a unit or layer, any number of nodes is provided. For example, 100 nodes are provided. Later or subsequent units may have more, fewer, or the same number of nodes. In general, for convolution, subsequent units have more abstraction. For example, the first unit provides features from the image, such as one node or feature being a line found in the image. The next unit combines lines, so that one of the nodes is a corner. The next unit may combine features (e.g., the corner and length of lines) from a previous unit so that the node provides a shape or building indication. For transposed-convolution to reconstruct, the level of abstraction reverses. Each unit or layer reduces the level of abstraction or compression.

[0078] The features of the nodes are learned by the machine using any building blocks. For example, auto-encoder (AE) or restricted Boltzmann machine (RBM) approaches are used. AE transforms data linearly, and then applies a non-linear rectification, like a sigmoid function. The objective function of AE is the expected mean square error between the input image and reconstructed images using the learned features. AE may be trained using stochastic gradient descent or other approach to learn, by the machine, the features leading to the best reconstruction. The objective function of RBM is an energy function. Exact computation of the likelihood term associated with RBM is intractable. Therefore, an approximate algorithm, such as contrastive-divergence based on k-step Gibb sampling or other, is used to train the RBM to reconstruct the image from features.

[0079] Training of AE or RBM is prone to over-fitting for high-dimensional input data. Sparsity or denoising techniques (e.g., sparse denoising AE (SDAE)) may be employed to constrain the freedom of parameters and force learning of interesting structures within the data. Enforcing sparsity within hidden layers (i.e., only a small number of units in hidden layers are activated at one time) may also regularize the network. In other embodiments, at least one unit is a convolution with ReLU activation or is a batch normalization with a ReLU activation followed by a convolution layer (BN+LeakyRU+convolution). Max pooling, upsampling, downsampling, and/or softmax layers or units may be used. Different units may be of the same or different type.

[0080] FIG. **7** shows a medical imaging system for image registration and/or therapy decision support. The system generates registered images on a display **500** to, for example, support therapy, diagnosis, and/or prognosis decisions.

[0081] The medical imaging system includes the display **500**, an image processor **502**, and memory **504**. The display **500**, the image processor **502**, and the memory **504** may be part of at least one medical imager **506**, a computer, a server, a workstation, or another system for image processing medical images from a scan of a patient. A workstation or computer without the medical imagers **506** may be used as the medical imaging system.

[0082] The medical imaging system shown in FIG. **7** includes two medical imagers **506***a*, **506***b*. In other embodiments, the medical imaging system includes more than two medical imagers **506**.

[0083] Additional, different, or fewer components may be provided. For example, a computer network is included for remote prediction based on locally captured scan data. As another example, a user input device (e.g., keyboard, buttons, sliders, dials, trackball, mouse, or other device) is provided for user interaction with the outcome prediction.

[0084] A first medical imager **506***a* is any number of different medical imaging devices including, for example, a magnetic resonance imaging (MRI) device, a computed tomography (CT) device, a positron emission tomography (PET) device, an ultrasound device, a dynaCT device, an angiogram device, and a mammography device. For example, the first medical imager **506***a* is an ultrasound device.

[0085] A second medical imager **506***b* is any number of different medical imaging devices including, for example, a magnetic resonance imaging (MRI) device, a computed tomography (CT) device, a positron emission tomography (PET) device, an ultrasound device, a dynaCT device, an angiogram device, and a mammography device. The second medical imager **506***b* is different than the first medical imager **506***a*. For example, the second medical imager **506***b* is an MRI device.

[0086] The medical imager **506** is configured by settings to scan a patient. The medical imager **506** is setup to perform